Purdue University
Purdue e-Pubs

GIS Day

GIS, Geoinformatics, and Remote Sensing at Purdue

11-17-2009

Validation of a Commercial Geographical Information Systems Database of Walking and Bicycling Destinations

Heather A. Whitcomb
Purdue University - Main Campus, hwhitcom@purdue.edu

Ellen K. Cromley
The Institute for Community Research

Kosuke Tamura
Purdue University - Main Campus, ktamura@purdue.edu

Steven J. Melly
Harvard School of Public Health

Sayali Kale
Purdue University - Main Campus, kales@purdue.edu

See next page for additional authors

Follow this and additional works at: http://docs.lib.purdue.edu/gisday

 Part of the Community Health and Preventive Medicine Commons, and the Urban Studies and Planning Commons

Whitcomb, Heather A.; Cromley, Ellen K.; Tamura, Kosuke; Melly, Steven J.; Kale, Sayali; Laden, Francine; James, Peter; Puett, Robin; Ben-Joseph, Eran; and Troped, Philip J., "Validation of a Commercial Geographical Information Systems Database of Walking and Bicycling Destinations" (2009). *GIS Day*. Paper 15.
http://docs.lib.purdue.edu/gisday/15

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Authors
Heather A. Whitcomb, Ellen K. Cromley, Kosuke Tamura, Steven J. Melly, Sayali Kale, Francine Laden, Peter James, Robin Puett, Eran Ben-Joseph, and Philip J. Troped

This article is available at Purdue e-Pubs: http://docs.lib.purdue.edu/gisday/15

Validation of a Commercial Geographic Information Systems Database of Walking and Bicycling Destinations

Heather A. Whitcomb[1], Ellen K. Cromley[2], Kosuke Tamura[1], Steven J. Melly[3], Sayali Kale[1], Francine Laden[3], Peter James[3], Robin Puett[4], Eran Ben-Joseph[5], Philip J. Troped[1]

[1]*Purdue University*, [2]*The Institute for Community Research*, [3]*Harvard School of Public Health*, [4]*University of South Carolina*, [5]*Massachusetts Institute of Technology*

Background

- Transdisciplinary research in public health, transportation, and urban planning has shown that destinations, such as stores, post offices, and physical activity facilities **within close proximity** to residences **are positively related** to recreational and transportation –related **physical activity**.

- Access to destinations has been measured several ways, (e.g., field audits and surveys individual's perceptions of their neighborhood).

- **Increasingly, researchers are using geographic information systems (GIS) and commercially available data on destinations to create objective measures of the built environment.**

- Recent studies have investigated the validity of GIS databases of physical activity facilities and food stores, but to our knowledge **less research has been conducted to validate a broader range of facilities that may serve as important walking and bicycling destinations.**

Objective

To **assess the validity of a commercially-available GIS database of facilities** that may serve as walking and bicycling destinations for adults.

Methods

- **12 census tracts** (4 high, 4 medium, and 4 low population density) were selected in two counties for analysis: **Hartford County,** Connecticut (Figure 1a) and **Tippecanoe County,** Indiana (Figure 1b).

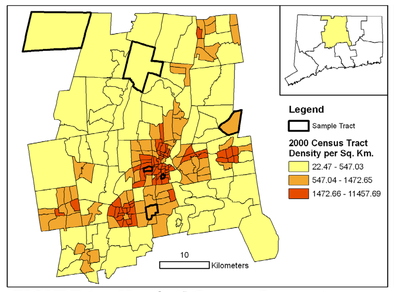

FIGURE 1a. Hartford County, Connecticut

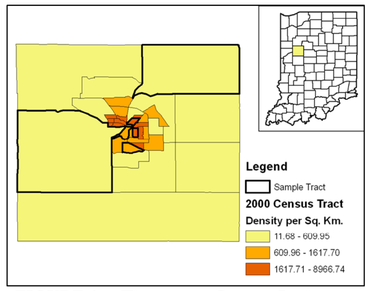

FIGURE 1b. Tippecanoe County, Indiana

- GIS database of facilities, current as of January 2007, was acquired by license.
- **402 facilities** in the database were categorized (by North American Industrial Classification System codes) into **5 destination types**: food and drink, social or cultural organizations, retail, services, and physical activity resources.
- **Field audits** to verify that the database agreed with field observations were conducted in Spring/Summer 2009.

- **Three levels of agreement** for comparisons between the database and the field were defined (Figure 2):
  1) **Perfect agreement:** located on the same street segment and had same proprietary name
  2) **Close to perfect agreement:** located on street segment and same type, but with different proprietary name
  3) **Adjacent street segment agreement:** located on adjacent street segment

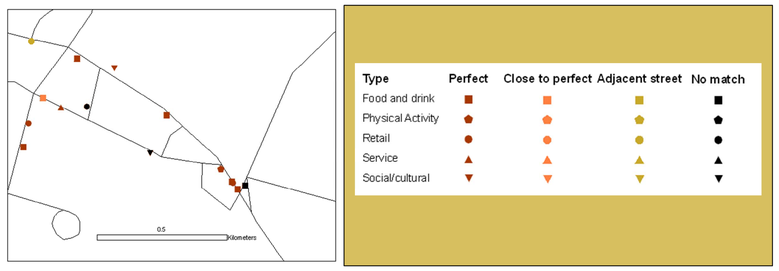

FIGURE 2. Example of destinations and observed levels of agreement

Statistical Analysis

- **Percent agreements** for destinations in the database that were found to be located in the field were calculated by level of agreement overall, and by county, population density, and destination type.
- **Chi-square** analyses were used to examine differences in percent agreement by county, population density, and destination type.

Results

- Overall, more than **two-thirds** of the 402 destinations in the database were found to be **in perfect agreement with field audits (Table 1)**. When all 3 levels of agreement were combined there was **86% agreement overall between the database and the field**.
- Chi-square analysis showed **no statistically significant differences** in agreement by county, population density, or destination type.

Table 1.
Agreement rates for database geocoded locations compared to field audit locations of destinations in 12 census tracts (*n* = number of destinations).

	Perfect	Perfect and Close to Perfect	Perfect, Close to Perfect, and Adjacent Street Segment
Total (*n* = 402)	67.7%	76.9%	85.8%
By County			
Tippecanoe County, IN (*n* = 200)	71.5%	80.0%	86.5%
Hartford County, CT (*n* = 202)	63.9%	73.8%	85.1%
By Census Tract Population Density			
High (*n* = 176)	71.0%	81.3%	88.1%
Medium (*n* = 122)	68.9%	75.4%	84.4%
Low (*n* = 104)	60.6%	71.2%	83.7%
By Type of Destination			
Food and drink (n = 139)	64.0%	75.5%	87.1%
Social/cultural (n = 115)	70.4%	77.4%	87.8%
Retail (n = 101)	67.3%	75.3%	81.2%
Service (n = 28)	64.3%	78.6%	85.7%
Physical activity (n = 19)	84.2%	89.5%	89.5%

Conclusions

- Results of this validation study demonstrated moderate to relatively good spatial accuracy of the commercial GIS database at the street-segment level with **more than 2/3 of the destinations correctly located in the field, overall**.

- Percent agreements generated in this study were **similar to** those in previous **studies of physical activity facilities and food stores** which were between 71%-73%.

- The commercially available GIS database in this study appeared to provide **a valid alternative to conducting extensive field audits or resident surveys.**

Acknowledgments

The study presented here is part of a larger investigation of the potential associations between the built environment and physical activity among women, and was funded through the National Cancer Institute (Grant#5-R21-CA125078-2).

The presenting author (Whitcomb) would like to thank the Society for the Advancement of Chicanos/Hispanics and Native Americans in Science for providing travel scholarship funding and the Alfred P. Sloan Foundation American Indian/Indigenous Graduate Scholarship Program for partially supporting her doctoral work.